Future Generation Computer Systems 28 (2012) 9–23

Contents lists available at SciVerse ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Scheduling efficiency of resource information aggregation in grid networks

P. Kokkinos *, E.A. Varvarigos

Department of Computer Engineering and Informatics, University of Patras, Greece
Research Academic Computer Technology Institute, Patras, Greece

ARTICLE INFO

Article history:
Received 1 March 2011
Received in revised form
23 June 2011
Accepted 30 June 2011
Available online 22 July 2011

Keywords:
Grid networks
Information aggregation
Scheduling
Stretch factor

ABSTRACT

We consider information aggregation as a method for reducing the information exchanged in a Grid network and used by the resource manager in order to make scheduling decisions. In this way, information is summarized across nodes and sensitive or detailed information can be kept private, while resources are still publicly available for use. We present a general framework for information aggregation, trying to identify issues that relate to aggregation in Grids. In this context, we describe a number of techniques, including single point and intra-domain aggregation, define appropriate grid-specific domination relations and operators for aggregating static and dynamic resource information, and discuss resource selection optimization functions. The quality of an aggregation scheme is measured both by its effects on the efficiency of the scheduler's decisions and also by the reduction it brings on the amount of resource information recorded, a tradeoff that we examine in detail. Simulation experiments demonstrate that the proposed schemes achieve significant information reduction, either in the amount of information exchanged, or in the frequency of the updates, while at the same time maintaining most of the value of the original information as expressed by a stretch factor metric we introduce.

1. Introduction

Grids consist of geographically distributed and heterogeneous computational and storage resources that may belong to different administrative domains, but are shared among users. The Grid resource manager (scheduler) receives user requests and assigns tasks to resources so as to optimize some objective function. Scheduling decisions are made based on static and dynamic resource information, including the computation and storage capacities, their availability, the number of tasks queued and other parameters of interest, which are usually collected by information services.

In this work, we study the operation of resource information aggregation in Grids. Through information aggregation, the resource characteristics are summarized before being sent to scheduling or other mechanisms (e.g., data managers), while Grid resources are still distributedly controlled. Resource-related information size and dynamicity grows rapidly with the size of the Grid, making the aggregation and use of this massive amount of information a challenge for the resource management system. In addition, as computation and storage tasks are conducted increasingly non-locally and with finer degrees of granularity, the flow of information among different systems and across multiple domains will increase. Information aggregation techniques are important in reducing the amount of information exchanged and the frequency of these exchanges, while at the same time maximizing its value to the Grid resource manager or to any other desired consumer of the information. An additional motivation for performing information aggregation is confidentiality and interoperability, since as more resources or domains of resources participate in the Grid, it is often desirable to keep sensitive and detailed resource information private, while resources are still being publicly available for use. For example, it will soon become necessary for the interoperability of the various cloud computing services (e.g., Amazon EC2 and S3, Microsoft Azure) that the large quantity of resource-related information is efficiently abstracted, before it is provided to the task scheduler. In this way, the task scheduler will be able to use efficiently and transparently the resources, without requiring services to publish in detail their resource characteristics. In any case, the key to information aggregation is the degree to which the summarized information helps the scheduler make efficient use of the resources, while coping with the dynamics of the Grid and the varying requirements of the users.

We propose several information aggregation techniques, which are presented in a general way, so as to permit their adaptation to specific situations, or their combination for the creation of new aggregation schemes. We are mainly interested in resource information that influences the decisions made by the Grid scheduler, while we do not consider parameters that cannot be directly aggregated (e.g., operating system version). Even

* Corresponding author at: Department of Computer Engineering and Informatics, University of Patras, Greece.
E-mail address: kokkinop@ceid.upatras.gr (P. Kokkinos).

0167-739X/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.future.2011.06.008

though the proposed information aggregation techniques are in line with the current hierarchical structure of Grid Networks and the related monitoring systems [1,2], our work follows a more theoretical approach, trying to identify generic issues that relate to aggregation, instead of studying how aggregation affects Grid Networks with specific characteristics. Specifically, we assume that resource sites are grouped into domains and the characteristics of the sites in each domain are aggregated before being sent to the resource scheduler. For aggregating the information, we use concepts and ideas derived from multi-criteria optimization [3]. In particular, each site is characterized by a vector of cost parameters that record its computation and storage capacity, their availability with time, and other parameters of interest, handling in this way the multi-dimensionality of the sites' characteristics. The cost vectors of the sites in a given domain are aggregated into a single (single point) or multiple (intra-domain aggregation) cost vectors for the entire domain, by performing appropriate associative operations to the site cost parameters. The domain cost vectors are collected from the monitoring system. We also introduce so-called *domination relations* that aim at reducing the number of vectors aggregated and stored. A resource site is said to be dominated by another one, when it is inferior to that with respect to all the resource parameters of interest. When a task request arrives, the scheduler applies an optimization function to the domain cost vectors in order to select the domain where the task will be executed. The task is then transferred to the selected domain and is assigned to a site in it, using exact resource information.

The undesirable side of information aggregation is that the efficiency of a scheduler using such information can be negatively affected. Even though there are obvious benefits in reducing through aggregation the amount of information exchanged and stored in a Grid, the job scheduler has less detailed and accurate information for decision making. This introduces an interesting *tradeoff between the amount and frequency of information exchanges and the value this information has for making efficient scheduling decisions*. We propose information aggregation schemes that produce aggregated (summarized) information of different quantity, granularity and therefore quality, improving or deteriorating scheduling decisions.

As a metric of the quality of the aggregated information we introduce the *Stretch Factor (SF)*, defined as the ratio of the task delay when the task is scheduled using complete resource information over the task delay when an aggregation scheme is used. We perform a large number of experiments to evaluate the proposed aggregation techniques, using a relatively simple simulation environment. Towards this end, we designed the simulations scenarios so that they are generic, evaluating various different scenarios: grids with many or few sites and domains, tasks with small or large workloads, resources with small or high computation capacities, etc. We also measure the number of resource information updates triggered by each aggregation scheme and the amount of information transferred. The simulation results show that the proposed schemes achieve significant information reduction, both in terms of size and of frequency of updates, while maintaining good scheduling quality. The uniformity or lack of uniformity of the sites' and tasks' characteristics is found to significantly affect the quality of the aggregation. In addition, we observe that that the type of parameters aggregated and the operators used for their aggregation play a significant role.

The present work handles information aggregation in Grids as a separate and important issue, attempting to identify the main issues, parameters, dependencies and side-effects related to the aggregation operation. Even though there are other works in Grids that consider aggregation as an available mechanism, they usually consider specific policies and scenarios, and they do not address through appropriate metrics the effect information aggregation has on the quality of the scheduling decisions made. In addition, other previous works focus on Data Networks where information aggregation is an old and well known mechanism used for making routing scalable and efficient. We use the same idea, making possible for a task scheduler to use transparently the globally available resources without knowing their full details, in the same way a router forwards a packet to the next gateway, without knowing all the intermediate nodes the packet will pass through, towards its destination.

The remainder of the paper is organized as follows. In Section 2 we report on previous work. The information aggregation problem is formulated in Section 3. Section 4 introduces the proposed aggregation techniques. In Section 5 we present an example of applying the techniques introduced in a hierarchical Grid Network. In Section 6 we experimentally evaluate the proposed techniques. Finally, conclusions are presented in Section 7.

2. Previous work

Information aggregation has been previously studied mainly in the context of hierarchical Data Networks [4], where it is performed on network-related parameters in order to facilitate routing. Resource information aggregation in Grids has not been studied in detail, despite its practical importance and its impact on the efficiency of the scheduling decisions. Actually, most scheduling algorithms proposed [5,6] make their decisions using exact resource information. In the following, we present prior work on the relevant topics.

Task scheduling is usually performed at two levels [7,8]; at the higher level, a central scheduler decides the site or domain a task will be executed on, while at the lower level a local scheduler selects the exact machine where the task will be executed. Most scheduling algorithms try to minimize the total average task delay [9] or maximize resource utilization, even though other performance metrics can also be used. The authors in [10] incorporate Grid economic models and propose scheduling algorithms that support deadline and budget constraints. Another criterion used in comparing scheduling schemes is the degree of fairness achieved among users in the use of the resources [11,12]. Finally, scheduling with advance reservations of computational and network resources has also received a great deal of attention in Grid and Data Networks [13,14]. A taxonomy of Grid resource schedulers is presented in [5,6]. In Grid Networks, resource information collection is performed by the monitoring systems. In [15] a number of monitoring systems are presented and categorized based on their architecture, as defined in the Global Grid Forum's Grid Monitoring Architecture (GMA) [16].

In existing Grid and Data Networks, sites/machines/nodes are organized in hierarchical structures (e.g., the EGEE [17] and the Internet). A node in a domain communicates with nodes belonging to other domains using specific border nodes. Hierarchical routing is a major issue for Data Networks, and is important for reducing the memory requirements at the routers (border nodes) for the very large topologies encountered in the Internet's infrastructure. A topology is broken down into several layers of hierarchy, thus downsizing the routing tables required, but this comes at the expense of an increase in the average path length. [18] is one of the first works investigating hierarchical routing, where clustering structures are introduced to minimize the routing tables required. Bounds are also derived on the maximum increase in the path length for a given table size. A central issue in hierarchical routing is topology information aggregation [4,19]. Aggregation techniques in hierarchical topologies try to summarize and compress the topology information advertised at higher levels. In order to perform routing and network resource allocation efficiently, the aggregated information should adequately represent the topology and the characteristics/metrics of the network.